A Lecture Series on DATA COMPRESSION

Lossy Compression — Transforms

Abdou Youssef (Speaker)

Anastase Nakassis (MDVTG Manager)

Motivation for Transforms

- Why transform the data
  - To decorrelate the data so that fast scalar (rather than slow vector) quantization can be used
  - To exploit better the characteristics of the human visual system (HVS) by separating the data into vision-sensitive parts and vision-insensitive parts
  - To compact most of the "energy" in a few coefficients, so that to discard most of the coefficients and thus achieve compression

Desirable Transforms

- Desirable Properties of transforms
  - Data Decorrelation, exploitation of HVS, and energy compaction
  - Data-Independence (same transform for all data)
  - Speed
  - Separability (for fast transform of multidimensional data)

- Various transforms achieve those properties to various extents
  - Fourier Transform
  - Discrete Cosine Transform (DCT)
  - Other Fourier-like transforms: Haar, Walsh, Hadamard
  - Wavelet transforms

- The Karhunen-Loeve Transform
  - Optimal w.r.t. data decorrelation and energy compaction
  - But it is data-dependent
  - And slow because the transform matrix has to be computed every time
  - Therefore, KL is only of theoretical interest to data compression

Different Perspectives of Transforms

- Statistical perspective
- Frequency perspective
- Vector space perspective
- End-use perspective (matrix formulation)

Matrix Formulation of Transforms

- Simply stated, a transform is a matrix multiplication of the input signal and the transform-matrix

- Each of the standard transforms mentioned earlier is defined by an $N \times N$ square non-singular matrix A_N

- Transform of a 1D discrete input signal (a column vector x of N components) is the computation of

$$y = A_N x$$

- Transform of an $N \times M$ image I is transform of each column followed by transform of each row. In matrix form, transform of image I is the computation of

$$J = A_N I A_M^t$$

- The inverse transform is simply $x = A_N^{-1} y$ for 1D signals, and $I = A_N^{-1} J (A_M^{-1})^t$ for images

Transform-Based Lossy Compression

- Compression of an image I:
  1. Transform I, yielding $J = A_N I A_M^t$
  2. Scalar-quantize J, yielding J'
  3. Losslessly compress J', yielding a bit stream B

- Image reconstruction
  1. Losslessly decompress B back to J'
  2. Dequantize J', yielding an approximation $\hat{J}$ of J
  3. Inverse-transform $\hat{J}$, yielding a reconstructed image

$$\hat{I} = A_N^{-1} \hat{J} {A_M^{-1}}^t.$$

Definition of The Matrices of the Standard Transforms

- Except for the case of the KL transform, the characterizing matrix A_N of each of the standard transforms is independent of the input, that is, A_N is the same for all 1D signals and images.

- In the following, the matrix A_N of each transform will be defined for arbitrary N, and then A_2, A_4 and A_8 will be shown

The Matrix of the Fourier Transform

- The matrix $A_N = (a_{kl})$ for the Fourier Transform:

$$a_{kl} = \sqrt{\frac{1}{N}} e^{-\frac{2\pi i}{N} kl}, \quad for\ k, l = 0, 1, \ldots, N-1$$

- Remark: $A_N^t = A_N = A_N^{-1}$

$$A_2 = \sqrt{\tfrac{1}{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$A_4 = \tfrac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

Let $a = \frac{\sqrt{2}}{2}(1+i)$ and $\overline{a} = \frac{\sqrt{2}}{2}(1-i)$

$$A_8 = \sqrt{\tfrac{1}{8}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \overline{a} & -i & -a & -1 & -\overline{a} & i & a \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ 1 & -a & i & \overline{a} & -1 & a & -i & -\overline{a} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -\overline{a} & -i & a & -1 & \overline{a} & i & -a \\ 1 & -i & -1 & -i & 1 & i & -1 & -i \\ 1 & a & i & -\overline{a} & -1 & -a & -i & \overline{a} \end{pmatrix}$$

The Matrix of the Discrete Cosine Transform (DCT)

- The matrix $A_N = (a_{kl})$ for DCT:

$$a_{0l} = \sqrt{\frac{1}{N}}, \quad for\ l = 0, 1, \ldots, N-1$$

$$a_{kl} = \sqrt{\frac{2}{N}} cos \frac{(l+\frac{1}{2})k\pi}{N}, \quad for\ 1 \leq k \leq N-1, 0 \leq l \leq N-1$$

- Remark: $A_N^{-1} = A_N^t$

$$A_4 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ \sqrt{1+\frac{\sqrt{2}}{2}} & \sqrt{1-\frac{\sqrt{2}}{2}} & -\sqrt{1-\frac{\sqrt{2}}{2}} & -\sqrt{1+\frac{\sqrt{2}}{2}} \\ 1 & -1 & -1 & 1 \\ \sqrt{1-\frac{\sqrt{2}}{2}} & -\sqrt{1+\frac{\sqrt{2}}{2}} & \sqrt{1+\frac{\sqrt{2}}{2}} & -\sqrt{1-\frac{\sqrt{2}}{2}} \end{pmatrix}$$

The Matrix of the Hadamard Transform

- The matrix $A_N = (a_{kl})$ for the Hadamard Transform is defined recursively:
  - $A_1 = (1)$
  - $A_N = \frac{1}{\sqrt{2}} \begin{bmatrix} A_{\frac{N}{2}} & A_{\frac{N}{2}} \\ A_{\frac{N}{2}} & -A_{\frac{N}{2}} \end{bmatrix}$

- Remark: $A_N^{-1} = A_N^t = A_N$

$$A_2 = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A_4 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

$$A_8 = \sqrt{\frac{1}{8}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}$$

The Matrix of the Walsh Transform

- The matrix A_N of the Walsh Transform is derived from the Hadamard matrix by permuting the rows of the latter in a certain way

- Remark: $A_N^{-1} = A_N^t = A_N$

$$A_2 = \sqrt{\tfrac{1}{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A_4 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

$$A_8 = \sqrt{\tfrac{1}{8}}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}$$

The Matrix of the Haar Transform

- The matrix $A_N = (a_{kl})$ for the Haar Transform, where $N = 2^n$:
  - $a_{0l} = \frac{1}{\sqrt{N}}$ for $l = 0, 1, \ldots, N-1$
  - for $k \geq 1$, $k = 2^p + q$, $0 \leq q \leq 2^p - 1$, $0 \leq p \leq n-1$

$$a_{kl} = \begin{cases} 2^{\frac{p-n}{2}} & \text{if } q2^{n-p} \leq l < (q+\frac{1}{2})2^{n-p} \\ -2^{\frac{p-n}{2}} & \text{if } (q+\frac{1}{2})2^{n-p} \leq l < (q+1)2^{n-p} \\ 0 & \text{otherwise} \end{cases}$$

- Remark: $A_N^{-1} = A_N^t$

$$A_2 = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad A_4 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{pmatrix}$$

$$A_8 = \sqrt{\frac{1}{8}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{pmatrix}$$

Vector Space Perspective

- Analog signals are treated as an infinite-dimensional functional vector space

- Finite Discrete are signals treated as finite-dimensional vector spaces

- In either case, the vector space has a basis $\{e_k \mid k = 0, 1, ...\}$

- A transform of a signal x is a linear decomposition of x along the basis $\{e_k\}$:
  - $x = \Sigma_k \, y_k e_k$, where the $\{y_k\}$ are real/complex numbers
  - Transform: $x \longrightarrow (y_k)_k$
  - $(y_k)_k$ is a representation of x

- Compression-related desirable properties of a vector-space basis
  - Correspondence with the human visual system
  - Specifically, only a very small number of basis vectors are relevant to (i.e., visible by) the HVS, while the majority of the basis vectors are invisible to the HVS
  - Uncorrelated decomposition-coefficients $(y_k)_k$

Relationship between the Vector Basis and the Matrix Formulation of Transforms

- Consider finite 1D discrete signals of N components
  - They form an N-dimensional vector space R^N
  - Any basis consists of N linearly independent column vectors $e_0, e_1, \ldots, e_{N-1}$
  - For any signal $x = (x_0\ x_1\ \ldots x_{N-1})^t$, $x = \Sigma_{k=0}^{N-1} y_k e_k$
  - That is, $\begin{pmatrix} x_0 \\ x_1 \\ \ldots \\ x_{N-1} \end{pmatrix} = (e_0\ e_1\ \ldots\ e_{N-1}) \begin{pmatrix} y_0 \\ y_1 \\ \ldots \\ y_{N-1} \end{pmatrix}$
  - Equivalently, $y = Ax$, where the columns of A^{-1} are the basis column vectors $e_0,\ e_1,\ \ldots\ ,\ e_{N-1}$

- Consider now $N \times M$ images
  - They form an NM-dimensional vector space $R^{N \times M}$
  - Any basis consists of $N \times M$ matrices $\{E_{kl}\}$ of dimensions $N \times M$
  - Following the same analysis as above, a transform $I \longrightarrow J = A_N I A_M^t$ corresponds to basis

    $$E_{kl} = (column\ k\ of\ A_N^{-1}).(column\ l\ of\ A_M^{-1})^t = e_k.e_l^t$$

    for $k = 0, 1, ..., N-1$ and $l = 0, 1, ..., M-1$
  - $I = \Sigma_{k,l} J_{kl} E_{kl}$

- Remark: For analog signals, the vector space is infinite-dimensional, and its basis is the infinite set of sine and cosine waves, to be addressed later

Visualization of DCT Basis Images

Fourier Transform

- Consider a function $x(t)$ that is either
  - of finite support $[0, T]$, or
  - periodic of period T

- Assume $x(t)$ to be square-integrable over $[0, T]$

- Fourier series of $x(t)$ is:

$$\sum_{k=0}^{\infty} a_k \cos \frac{2\pi}{T}kt + \sum_{k=0}^{\infty} b_k \sin \frac{2\pi}{T}kt, \ \ or$$

- In (more elegant) complex form:

$$\sum_{k=-\infty}^{\infty} y_k e^{\frac{2\pi}{T}ikt}, \qquad where\ y_k = \frac{1}{T}\int_0^T x(t)e^{-\frac{2\pi}{T}ikt}dt$$

- Fourier Transform: $x(t) \longrightarrow (y_k)_k$

- Theorem: For all natural signals,
  - $\frac{x(t_+)+x(t_-)}{2} = \Sigma_{k=-\infty}^{\infty} y_k e^{\frac{2\pi}{T}ikt}$
  - If x is continuous at t, $x(t) = \Sigma_{k=-\infty}^{\infty} y_k e^{\frac{2\pi}{T}ikt}$

- Therefore, $x(t)$ is largely representable by $(y_k)_k$

- Theorem: $y_k \longrightarrow 0$ as $|k| \longrightarrow \infty$

- The representation $x(t) = \Sigma_{k=-\infty}^{\infty} y_k e^{\frac{2\pi}{T}ikt}$ is periodic of period T. Thus, even if $x(t)$ is defined over $[0, T]$ only, the Fourier series "periodizes" $x(t)$

Connection with The Human Visual System

- $e^{\frac{2\pi}{T}ikt}$ is periodic of period $\frac{T}{k}$; thus, its frequency is $\frac{k}{T}$
- The higher k, the higher the frequency
- In $x(t) = \Sigma_{k=-\infty}^{\infty} y_k e^{\frac{2\pi}{T}ikt}$, y_k is the k-th frequency content of $x(t)$
- Experiments have shown that
  - suppressing a y_k (along with y_{-k}) for any high frequency k causes HARDLY VISIBLE or NO VISIBLE change to $x(t)$
  - suppressing a y_k (along with y_{-k}) for some low frequency k causes VISIBLE changes to $x(t)$
- Thus, the HVS is sensitive to low-frequency data but insensitive to high-frequency data

Connection to Compression

("First Cut")

- Facts
  - $x(t) = \Sigma_{k=-\infty}^{\infty} y_k e^{\frac{2\pi}{T}ikt}$
  - $y_k \longrightarrow 0$ as $|k| \longrightarrow \infty$

- $\hat{x}(t) = \Sigma_{k=-r}^{r} y_k e^{\frac{2\pi}{T}ikt}$ is a good mathematical <u>and</u> visual approximation of $x(t)$

- The faster the decay of y_k, the smaller r can be

- Thus, $(y_k)_{|k| \leq r}$ is a very small representation of x ($\hat{x}$ to be precise), leading to high compression

Treatment of Discrete Signals (Discrete Fourier Transform)

- Sample N values (x_l) of $x(t)$ at N points $l\frac{T}{N}$, that is, $x_l = x(l\frac{T}{N})$ for $l = 0, 1, ..., N-1$.

- $x_l = x(l\frac{T}{N}) = \Sigma_k\, y_k e^{\frac{2\pi}{T} ikl\frac{T}{N}} = \Sigma_k\, y_k e^{\frac{2\pi}{N} ikl}$

- Since (x_l) is discrete and finite, there is no need to keep an infinity of y_k's; rather, $y_0, y_1, ..., y_{N-1}$ are sufficient. That is,

$$x_l = \sum_{k=0}^{N-1} y_k e^{\frac{2\pi}{N} ikl}, \quad l = 0, 1, ..., N-1$$
$$y_k = \sum_{l=0}^{N-1} x_k e^{\frac{2\pi}{N} ikl}, \quad k = 0, 1, ..., N-1$$

- DFT: $(x_l)_l \longrightarrow (y_k)_k$

- Put in matrix form: $y = A_N x$, where $A_N = (e^{\frac{2\pi}{N} ikl})_{kl}$

- Again, for large k, y_k can be suppressed is broadly quantized

Why Use DCT rather than DFT (Boundary Problems of the Fourier Transforms)

- Discontinuities at the boundaries cause large high-frequency contents

- Eliminating those frequency contents cause boundary artifacts (known as Gibbs phenomenon, ringing, echoing, etc.)

- Consider the function

$$x(t) = \begin{cases} \frac{t}{T} & \text{if } 0 \leq t \leq \frac{T}{2} \\ (2\epsilon - 1)\frac{t}{T} - \epsilon + 1 & \text{if } \frac{T}{2} \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$

- Its Fourier series is:

$$x(t) = \frac{\epsilon + 1}{4} + \sum_{k \neq 0} \left[\frac{(1-\epsilon)\left((-1)^k - 1\right)}{2(\pi k)^2} + \frac{\epsilon}{2\pi k} i \right] e^{\frac{2\pi i}{T} kt}$$

- Special case $\epsilon = 0$:

$$x(t) = \frac{1}{4} + \sum_{k \neq 0} \frac{(-1)^k - 1}{2(\pi k)^2} e^{\frac{2\pi i}{T} kt}$$

- Special case $\epsilon = 1$:

$$x(t) = \frac{1}{2} + \sum_{k \neq 0} \frac{i}{2\pi k} e^{\frac{2\pi i}{T} kt}$$

Relation of DCT to FFT

- Let (x_l) be an original signal, and (y_k) its DCT transform, $l = 0, 1, ..., N-1$

- Shuffle x to become almost symmetric; that is, create a new signal (x'_l) by taking the even-indexed terms followed by the reverse of the odd-indexed terms:
  - $x'_l = x_{2l}$ and $x'_{N-l-1} = x_{2l+1}$, for $0 \le l \le N/2 - 1$

- y'=DFT(x');

- $y_0 = \sqrt{\frac{1}{N}} Real(x'_0)$,
  and $y_k = \sqrt{\frac{2}{N}} Real(e^{\frac{-\pi k}{2N}} x'_k)$, $k = 1, 2, ..., N-1$

Quantization in DCT-based Compression

DCT vs. KL

- For most natural signals, the KL basis and the DCT basis are almost identical

- Therefore, DCT is near optimal (in decorrelation, energy compaction, and rme distortion) because KL is optimal

- Unlike KL, DCT is not signal-dependent

- Hence the popularity of DCT